The Indonesian Journal of Geography, Vol.30, No.75
June 1998, pp. 33-42

THE EFFECT OF GEOMETRICALLY CORRECTED LANDSAT TM DATA ON THE ACCURACY OF THE LANDUSE CLASSIFICATION RESULTS

A Case Study in Semarang Municipality and Its Vicinity

by
Bambang Sulistyo and Hartono*

ABSTRACT

This research was aimed at first, studying and comparing between the accuracies of the geometrically and non-geometrically corrected digital Landsat data; and second, evaluating these classification accuracies to ascertain the possibility of using them as an input into the process of building up a Geographical Information System.

The method used was digital landuse classification by applying the data's maximum likelihood algorithm in the two approaches. The first approach involved classification prior to geometric correction (reference) and the second classification after geometric correction (known as transformed result). Analysis was then carried out through the overlay technique between the first (reference) and the second (transformed) results. From the process of analysis was found an error matrix depicting individual and overall accuracies; and omisions and commission of errors.

The result shows that the overall accuracy of landuse classification after geometric correction is > 80%, and this implies that such a result can be used as an input into GIS. However, this overall accuracy varied according to the technique applied; i.e, nearest neighbour interpolation (88,73%), cubic convolution (83,69%) and bilinear interpolation (82,67%). But the choice of which technique to be used depends on the average increase and decrease in area. As such, the use of the nearest neighbour interpolation, bilinear interpolation and cubic convolution techniques resulted into an areal increase of 19.54%, 24.80%, and 24.93%, and a reduction of 17.17%, 24.60%, and 27.87% respectively.

---

* Ir. Bambang Sulistyo, Dipl.IGP., M.Si is a teaching staff member at the Faculty of Agriculture. Bengkulu University, Bengkulu, Indonesia and Dr. Hartono, DESS is a teaching staff member as well as the Head of PUSPIC, Faculty of Geography, Gadjah Mada University, Yogyakarta, Indonesia.

INTRODUCTION

Development activities in all sectors have to be supported by a good planning process. A vital prerequisite for this process in the availability of complete, detailed and up-to-date data in the form of maps, is usually obtained through the implemantion of resource inventory where in, the amount, type, distribution, and quality of terrestially related resources can be presented (Sutanto, 1987).

However, the acquisition of this up-to-date base map is not an easy task even with the use of conventional technology because it requires more time, energy and funds. This can be examplified by the fact that the available base maps scales 1 : 50,000 covering Indonesia's land do only cover about 70% of the total land (Sutanto, et al.,1996).

Remote Sensing is a relatively new technology offering a solution to the preparation of up-to-date maps at a relatively low cost in term of time, energy, and finance that can enable wider areal coverage. Since remote sensing data can be in a digital form, extraction of information that can be presented on to an up-to-date map vitalises data processing.

For comprehensive planning purposes, the results from processed remotely sensed digital data are used as an input into GIS (Belward & Valenzuela, 1990). These data are usually obtained as the resultants of landuse classification process. However, the accuracy of the results out of this classification process calls for a comprehensive mastery in order to decide whether it could be used as an input into GIS or not.

One factor that greatly affects the results of landuse classification is that of the geometric error, whose elimination is known as geometric correction. In other word, this geometric correction is a technique of transforming remotely sensed data to enable it contain all map characteristics under a given scale and projection (Mather, 1987). Geometric correction involves intensity (spectral) interpolation i.e., the determination of a new brightness value at a newly positioned pixel through the process of resampling.

The accuracy limit of a classification result that appropriates it as an input into GIS varies amongst researchers. Anderson et al., (1976) recommended that the accuracy of landcover/landuse classification derived from landsat MSS image at Level 1 should be at least 85%, whereas Daels and Antrop (1981) stated it as at least 80% (in Gunawan, 1993). On the other hand, Justice (1978 in Townshed, 1981) distinguishes the accuracy of classification results into three categories namely: Good (>85%), Moderate (70-85%), and Poor (<70%).

This research is based on the question: how do the accuracies of geometrically and non-geometrically corrected digital classification results differ ? This question is attributed to the fact that many researchers have suggested that process of classification has to be executed prior of geometric correction to ensure that the originality of the data is maintained, i.e. the resampling process does not distort the data (Mather, 1987; Jensen, 1986; Campbell, 1987). But sensors and resolutions, geometric correction is necessitated before classification (Chavez and Bowell, 1988). As such there are cases which require geometric correction before classification, for instance when areal sampling identification is done using map co-ordinates, and also when working with multi-images (Thomas et al., 1987).

Digital classification is a process by which data are grouped together basing on spectral analysis and then finally calssified as per their similarity characteristics of the pixel values (Lillesand and Kiefer, 1995). This classification is based on the recognition of spectral patterns though the application of the parallelepiped, minimun distance to the mean, and maximum likelihood techniques, etc. The best technique is the best that can yield a good accuracy level because it takes into account the variances and correlations between the chosen samples.

In relation to the need of geometric correction, supervised digital classification was implemented in two approaches. The first approach involved classification prior to geometric correction, and the second one after geometric correction as shown in figure 1.

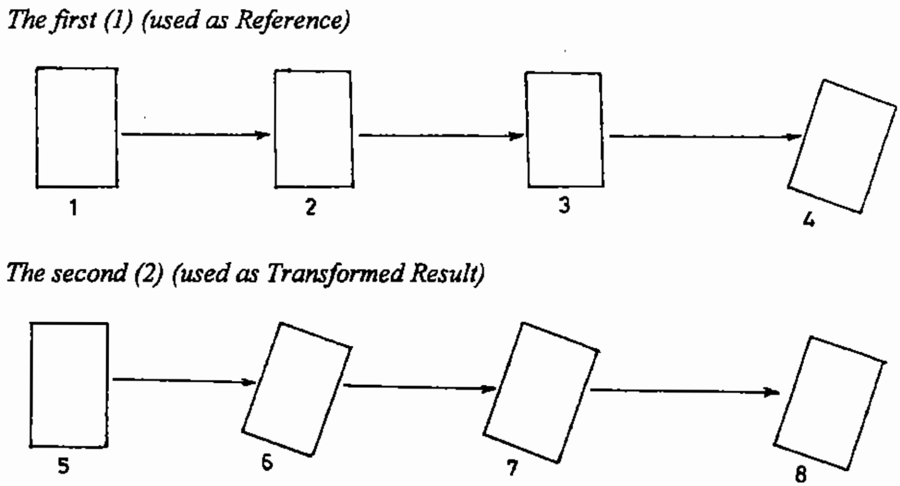

Figure 1. Two approaches for geometric correction (Bambang Sulistyo, 1997)

Where :

1. Non-geometric correction images
2. Sample identification with same area and boundary in 1
3. Results of digital classification
4. Final result after performing geometric correction
5. Non-geometric correction images
6. Geometric correction images
7. Sample identification with same area and boundary in (6) as in (1)
8. Final results of digital classification

When:
1. landcover/landuse sample identification with a certain area and boundary in (1) is also applied to (2),
2. both approaches are using Maximum Likelihood classifier and
3. the first digital classification results are used as a reference, and so it will result into a difference in pixel numbers between the first and second classification results because the statistical values are changed due to the resampling process which during geometric correction involves interpolation techniques such as either nearest neighbour, bilinear, or cubic convolution.

The application of overlay analysis enables the comparison of both kinds of results (first and second) and the ultimated result is an error matrix wich in turn depicts the relationship between the reference (first) and transformed (second) results. This error matrix is then stated as either an overall accuracy, individual accuracy, or as an ommision and/or a commision of error.

To establish the possibility of adopting the second classification result (transformed) as an input into GIS, threshold values have to be set too. Generally for Indonesia, and most especially for the study area, this result is appropriated when the landuse/landcover classification accuracy is greater than or equal to 80% as recommended by both Daels and Antrop. This limit established noticing that:
- digital proses of remotely sensed data is relatively a new technique that has not yet been widely and routinely used as compared to the analog methods like aerial photograph,
- landuse/landcover situation in Indonesia varies in term of both type and size as opposed to the situation prevalent in either USA or Europe that is more regular.

METHODOLOGY

The study area covers part of Semarang municipality and Kendal regency, Central Java Province (Figure 2). This research made use of both a topographic map, scaled 1 : 50,000 (Ungaran, Kendal, Sukorejo, and Semarang sheets); and four bands (2,3,4, and 5) of Landsat TM Satellite digital data, path/row 128/065 acquired in June, 1994. Analysis was conducted by using ILWIS (Integrated Land and Watershed Information System), version 1.4, a PC based software designed by ITC, The Nehterlands.

The procedure of the research involved the following steps: preparation, digital classification (interpretation), field work, digital reclassification (reinterpretation) and overlay analysis; and finally report making, in that order.

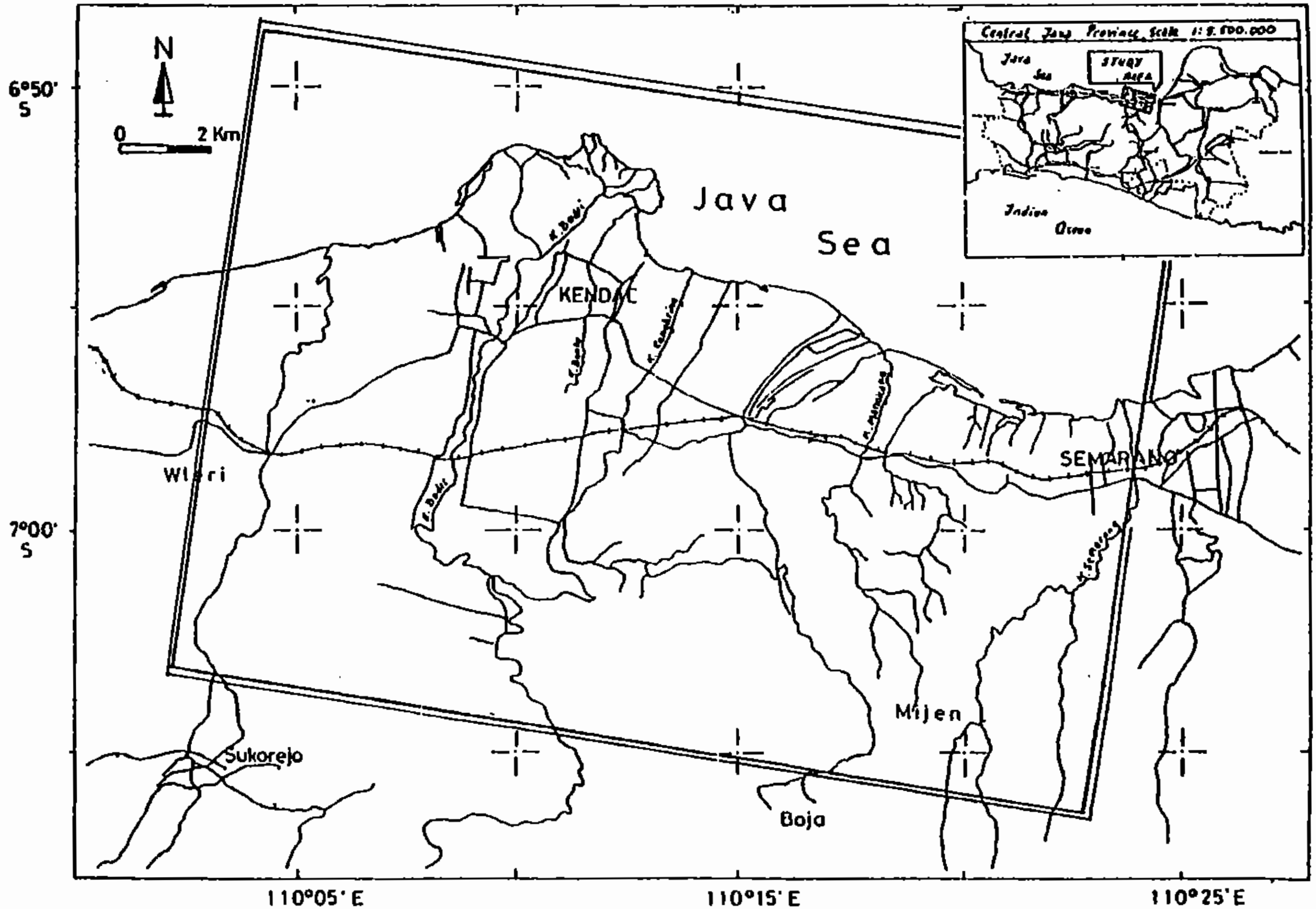

Figure 2. Location of the study area.

Preparation

This step was basically a literature review and the search for Landsat TM data and maps as well as any other data that would be deemed useful.

Digital Classification/Interpretation

In this step, supervised digital landuse/landcover classification were applied with the help of the maximum likelihood algorithm technique both before and after geometric correction.

Fieldwork

This step was intended for the purpose of verification, i.e ground truthing the digital classification results in relation to the actual landuse/landcover. Some sample areas were observed for their features in term of landuse/landcover, soil type, crown closure density, and environmental condition.

Digital Reclassification/Reinterpretation

Reinterpretation and overlay analysis step are aimed at improving upon the classification results in case of any error during the preliminary classification step. Analysis by overlaying the second result on the first was done to obtain the error matrix there by depicting the relocationship between them.

Report Writing

This step which involves the summarisation of the research results is important for purposes of presentation and information storing. The results from the processing and analysis of the research data were written in a report right from the start to the end or the final step.

RESULTS AND DISCUSSION

The results of digital landuse classification

The obtained results of digital landuse classification and their pixel numbers are presented in table 1.

Table 1. The results of digital Landuse classification and the corresponding pixel numbers

Code Landuse	Transformed Results			Master
	HTNN Pixel nbr	HTBI Pixel nbr	HTCC Pixel nbr	Pixel nbr
01 Wetland Rice	191908	189116	191426	184960
02 Tobacco Plantation	24313	18031	20372	23292
03 Teak Forest	207269	213092	211672	216875
04 Open Sea	319630	317450	318664	314917
05 Coastal Fish Pond	52783	51700	52424	59075
06 Sugarcane Plantation	20050	17662	18520	17978
07Field Crops	31169	30922	31639	29477
08 Shrubs	3938	3196	3762	3704
09 Mixed-Garden/Kampong	141770	132007	137104	158214
10 Grass	19614	18172	17776	20093
11 Open & Built up Land	25257	23283	24392	24990
12 Rubber Estates	23044	16013	18907	22598
13 Other	475283	505384	489370	449015

Notes :

1. HTNN : Transformed images from nearest neighbour interpolation technique
2. HTBI : Transformed images from bilinear interpolation technique
3. HTCC : Transformed images from cubic convolution interpolation technique
4. Factor 900 was used to convert pixel numbers to m^2

Comparison between the First (Reference) and Second (Transformed) Landuse Classification Results

By analytically overlaying the results of the first and the second landuse/cover classification, it was possible to evaluate the difference between them as depicted by the error matrices expressed in term of their individual accuracies, omision and commission of errors for every class, as show in Table 2.

DISCUSSION

The Effect of Geometric Correction on the Landuse Classification Results

Table 2 shows that HTNN image resulted into an overall accuracy of 88.73% that was closest to the reference compared to HTBI (82.67%), or to HTCC (83.69%). This is attributed to the fact that the nearest neighbour interpolation technique did not bad to new pixel values. In other words, the pixel values remained relatively constant and so the results of landuse classification either before or after geometric correction remained relatively similar. The difference of about 11.27% realized is attributed to the effect of pixel repetition as an inherent characteristic of the nearest neighbour interpolation technique which are exemplified by the presence of discontinuation lines in the images.

Table 2. Comparison of the Results of the First and the Second Landuse Classifications (%)

No Landuse	HTNN			HTBI			HTCC		
	Omis	Commi	Indiv	Omms	Commi	Indiv	Omms	Commi	Indiv
01 Wetland Rice	4.53	9.47	90.98	12.52	16.35	84.25	16.02	16.02	84.76
02 Tobacco Plantation	10.44	15.78	85.02	40.78	17.42	77.27	21.38	21.38	75.23
03 Teak Forest	24.00	21.62	77.86	27.80	29.28	17.15	28.08	28.08	71.99
04 Open Sea	0.22	0.56	99.44	1.00	0.50	99.50	0.58	0.58	99.42
05 Coastal Fish Pond	2.75	1.01	98.97	8.25	4.10	95.81	4.47	4.47	95.42
06 Sugarcane Plantation	4.70	14.68	86.65	18.08	16.76	83.02	18.89	18.89	81.61
07 Field Crops	16.82	20.08	80.55	28.72	33.42	68.08	33.99	33.99	67.98
08 Bushes	54.39	67.48	40.33	59.52	65.15	38.32	71.79	71.79	36.64
09Garden/Kampong	18.81	6.73	92.34	29.27	7.45	90.48	9.13	9.13	88.97
10 Grass	48.52	44.53	53.62	54.86	44.35	50.44	39.67	39.67	53.57
11 Open and built land	9.08	7.85	92.05	25.17	15.11	83.20	17.33	17.33	81.71
12 Rubber Estates	13.54	13.90	86.15	37.80	7.44	89.32	12.15	12.15	85.42
13 Others	15.45	30.32	73.61	18.52	65.12	55.58	50.91	50.91	60.98
Average	17.17	19.54		27.87	224.80		24.60	24.93	
Overall Accuracy		88.73			82.67			83.69	
Accuracy Reduction		11.27			17.33			16.31	

Notes :

1. Omis is Omission of Error
2. Commi is Commission of Error
3. Indiv is Individual Accuracy

* Mixed Gardens

Mather (1987) stated that if sample areas are more homogeneous, the bilinear interpolation technique usually yields a smaller standard deviation compared to that of the cubic convolution technique. In cubic convolution, smoothing was done based on the consideration of 16 surrounding pixels by use of a polynomial function, while bilinear interpolation based on consideration of only 4 surrounding pixels used linear functioning, and the nearest neighbour interpolation technique considers only the closest neighbouring pixels.

Relatively homogenous sample areas resulted into more unclassified pixels are indicated by the classification result of HTCC which is closest to the result of reference, i.e, differing by 16.31% compared to HBTI landuse classification result which differs from the reference result by 17.33%.

From this results, it can be concluded that landuse/landcover classification results after performing geometric correction, whether by use of the nearest neighbour, or bilinear interpolation or cubic convolution technique are still greater then the set threshold value of 80%. This implies that these classification results can be used as an input into the process of developing Geographical Information Systems for planning activity by using GIS procedure.

The Effect of the Results of Landuse Classification after Geometric Correction on the Landuse Area

As a consequence of the interpolation technique and hence resampling processes used during geometric correction, samples statistical values were altered by the difference between the pixel numbers of the reference classification (before geometric correction) and the transformed result (after geometric correction).

The changes in area are denoted by the degrees of omission and commision of error. The former shows a decrease in area while the latter an increase. From table 2 we realize that the three interpolation techniques yielded varying ranges of incremental effects on area, i.e, nearest neighbour (0.56% - 67.48%), bilinear (0.50% - 64.15%), and cubic convolution (0.58% - 71.79%); whereas their areal reduction effects varied too at the ranges of (0.22% - 54.39%), (1.00 - 59.52%), and (0,73% - 54.23%) respectively.

Table 2 also indicates that the technique with the least average incremental effect was the nearest neighbour interpolation technique (19.54%), followed by the bilinear interpolation (24.80%), and finally the cubic convolution (24.73%). However the technique with the greatest reduction effect on average was bilinear interpolation (27.87%) followed by the cubic convolution (24.60%) and finally nearest neighbour technique (17.17%) as the least.

CONCLUSIONS

Landuse classification after performing a geometric correction has an overall accuracy greater then the threshold value of 80% and therefore, it implies that classification results after geometric correction can be used as an input into GIS. Also since the overall accuracies of these three methods do differ, i.e, nearest neighbour (88.73%), cubic convolution (83.69%) and bilinear interpolation (82.67%), it means that the best technique for this landuse classification is the nearest neighbour interpolation technique for its overall accuracy is highest.

By closing to use the nearest neighbour interpolation technique, a least incremental effect on area, i.e, averagely 19.54% was realized as opposed to the other two approaches, billinear interpolation (24.80%) and cubic convolution (24.93%). The same can also be seen in their reduction effects of area which are in the magnitudes of 17.17%, 24.60%, and 27.87% respectively.

ACKNOWLEDGEMENT

The authors are indebted to Prof.Dr. Sutanto, a teaching staff member in the Faculty of Geography, Gadjah Mada University, Yogyakarta, who whole heartedly availed us with Landsat thematic mapper digital data.

REFERENCES

Anderson, J.R., Hardy. E.E., Roach, J.T., and Witmer, R.E., 1976, *A Landuse and land Cover Classification System for Use with Remote Sensor Data*, United States Government Printing Office, Whashington.

Mambang Sulistyo, 1997, Evaluation Ketelitian Hasil Klasifikasi Penggunaan Lahan dari Data Landsat TM Setelah Koreksi Geometris (Studi Kasus Kotamadya Semarang dan Sekitarnya), *Tesis*, Program Studi Penginderaan Jauh, Program Pasca Sarjana, UGM, Yogyakarta.

Belward, S.S., and Valenzuela, C.R., 1990, *Remote Sensing and Geographical Information Systems for Resource Management in Developing Countries*, Kluwer Academic Publishers, London.

Campbell, J.B., 1987, *Introduction to Remote Sensing*, Guilford Press, New York.

Gunawan, T., 1993, Penginderaan Jauh Terapan Untuk Studi Pedesaan, *Bahan Kuliah Pada Program Studi Penginderaan Jauh*, Program Pasca Sarjana, Universitas Gadjah Mada, Yogyakarta.

Jensen, J.R., 1986, *Introductory Digital Image, A Remote Sensing Perspective*, Prentice Hall, Inc., New York.

Lillesand, T.M., dan Kiefer, R.W., 1995, *Remote Sensing and Image Interpretation*, John Wiley and Sons, New York.

Mather, P.M., 1987, *Computer Processing of Remotely Sensed Images, An Introduction*, John Wiley and Sons, New York.

Sutanto, 1987, Pengalaman Penggunaan Citra SIR-A dan SIR-B Wilayah Indonesia, *Makalah pada Kursus Radar Penginderaan Jauh dalam rangka Kerjasama Penginderaan Jauh Indonesia-Australia di BAKOSURTANAL*, Bogor.

Sutanto, Hartono, Danoedoro, P., 1996, Pengembangan Ilmu Pengetahuan dan Teknologi Pemetaan Tanah Serta Sumberdaya Manusia Pemetaan Tanah, *Seminar Nasional Pemetaan Tanah I*, Jakarta.

Thomas, J.L., Benny, V.M., and Ching, N.P., 1987, *Classification of Remotely Sensed Images*, Adam Hilger, Bristol, England.

Townshend, J.R.G., 1981, *Terrain Analysis and Remote Sensing*, George Allen & Unwin London.

Valenzuela, C.R., 1991, Basic principles of Geographic Information System, *In Remote Sensing and GIS for Resources Management in Developing Countries*, The Nehterlands.